DOCUMENT RESUME

ED 442 042 CG 030 094

AUTHOR	Fennerty, Dan; Lambert, Connie; Majsterek, David
TITLE	Behavior Rating Scales: An Analysis.
PUB DATE	2000-00-00
NOTE	15p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Behavior Disorders; *Behavior Rating Scales; Child Development; Elementary School Students; Elementary Secondary Education; Labeling (of Persons); Personality Assessment; Secondary School Students; Special Education; Student Behavior; Student Evaluation
IDENTIFIERS	Behavior Rating Profile; Child Behavior Checklist; Conners Rating Scales; Walker McConnell Scale of Soc Compet and Sch Adj

ABSTRACT

In order to qualify for special education services, students must meet criteria that signal a serious behavior disorder. The Executive Committee of the Council for Children with Behavior Disorders (1998) expressed the need to move away from norm-referenced measures and recommended utilizing a more functional assessment approach. Members of the counseling field also are advocating for behavioral assessment alternatives to more formal procedures. In light of these recommendations, this paper reviews four widely utilized behavior rating scales (Behavior Rating Profile, Child Behavior Checklist System, Conners' Rating Scales, Walker/McConnell Scales of Social Competence and School Adjustment) regarding their recommended uses and behavioral descriptors. It suggests the following recommendations when considering a rating scale: (1) they should be supplemented with directly observed behavioral data; (2) educators should be aware that scales reflect perceptions about students and multiple informants and inter-rater reliability checks can corroborate or contradict these perceptions; (3) before using a rating scale, educators should make sure it reflects overall goals of the assessment process; (4) care should be taken so that information about the student is not skewed toward the negative; and (5) data on informant perceptions should be carefully considered. (Contains 1 table and 13 references.) (JDM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 442 042

Behavior Rating Scales: An Analysis

By

Dan Fennerty, Connie Lambert, and
David Majsterek

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Fennerty

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

CG030094

BEST COPY AVAILABLE

CG

Behavior rating scales: An analysis

by Dan Fennerty, Connie Lambert, and David Majsterek

Central Washington University

Abstract. Four widely utilized behavior rating scales were reviewed regarding recommended uses and behavioral descriptors. Two independent reviewers reviewed the items and inter-rater reliability was calculated. Recommendations for use of behavior rating scales in an overall assessment process are provided.

introduction.

Students with behavior disorders (BD) manifest wide-ranging problems in academic, literacy, social-emotional, motivational, and cognitive realms (Swicegood, 1994). In order to qualify for special education services, students must meet criteria that signal a serious behavior disorder. Inappropriate placement and programming decisions concerning these students are sometimes due to over reliance on standardized measures (Hallahan & Kauffman, 1991). The Executive Committee of the Council for Children with Behavior Disorders (CCBD) (1989) express the need to move away from these norm-referenced measures, and recommend utilizing a more functional assessment approach involving direct observation of students' learning environments, curriculum-based measures, and student work products.

In short, the field seems to advocate for behavioral assessment alternatives to more formal procedures. Mash and Terdal (1981) define behavioral assessment as "... a

CG 030094

range of deliberate problem-solving strategies for understanding children and childhood disorders" (p. 8). Further, these assessments should measure target behaviors over time (Mash & Terdal, 1981). The authors go on to say that the outcome of such problem-solving strategies should be a picture of the child that is detailed, accurate, and useful for planning appropriate interventions.

Shapiro (1987) suggests that an advantage of direct observation behavioral assessment practices is that such procedures can be used not only for diagnosis but to develop and identify strategies that would best address student needs (e.g., target behaviors) more effectively. He points out that the purpose of behavioral assessment is to provide information pertaining directly to the development of instructional programs.

Diagnosis, design, and evaluation are identified as purposes of behavioral assessment (Mash & Terdal, 1981). Diagnosis and design involve gathering information relevant to the development of effective interventions. Evaluative assessment is used to determine whether (1) objectives are being met; (2) changes are related to interventions; (3) changes are long-lasting; and (4) interventions and proposed outcomes are acceptable to students, teachers, and parents (Mash & Terdal, 1981).

Behavioral assessments are used to measure behavioral, social and emotional difficulties of children and youth (Merrell, 1994). Relevant information is typically collected via informants. Data are gathered using a checklist or rating scale format to which the evaluator responds "not true at all" through "very much true." For example, the Conners Teacher Rating Scales presents items (e.g., "cries often and easily") to which the evaluator responds by using a four point scale: zero indicating "not at all," one indicating "just a little," two indicating "pretty much", and three indicating "very much"

(Conners, 1990). On such published instruments, behavioral descriptors are typically worded either positively (e.g., "I am happy") or negatively (e.g., "Does not obey unless threatened with punishment"). Scores are compiled and compared to a standardization sample. Four informant-type behavior rating scales were reviewed for this study and will be discussed.

Behavior Rating Scales

Merrell (1994) defines behavioral rating scales as "...a standardized format for the development of summative judgments about a child or adolescent's behavioral characteristics, supplied by an informant [usually a parent or teacher] who knows the subject well" (p. 66). Merrell (1994) goes on to say that "behavior rating scales are less direct than either direct observation or structured behavioral interviewing in that they measure *perceptions* of specified behaviors rather than provide a firsthand measure of the existence of the behavior" (p. 66). Scores from behavior rating scales are compared to those of a normative sample to help determine if the identified behaviors are within "normal" limits (Bellack & Hersen, 1988). Information from the rating scale can be used to target specific behaviors that warrant further attention.

Due to certain aspects of behavior rating scales, Witt, Elliott, Daly, Gresham, and Kramer (1998) note the following:

1. Rating scales presume that the evaluator has made sufficient observations and can reliably note the relative frequency of specific behaviors; since the precision of rating scales is relative, they should be supplemented with more direct methods.

2. Ratings of behavior are evaluative judgments affected by the environment and a rater's standards for behavior; consequently, behavior may change from one setting to another. Therefore the use of multiple informants in multiple environments is recommended.
3. The social validity of behaviors should be understood in context since behavior in a specific situation may reflect the tolerance levels and/or standards of the rater; again the use of multiple informants in multiple environments is recommended.
4. Multiple raters of the child's behavior may agree only moderately and different raters may perceive behavior differently. Therefore, informant perceptions should be corroborated through inter-rater reliability data on perceptions of student performance.

Despite the potential limitations, behavior rating scales can provide useful, economical, reliable, and valid information about a person's social skills and behavior (Witt, et. al, 1998). This potential is increasingly realized as variety increases. Multiple informants is one way to increase the opportunity that a picture of the subject is detailed, accurate, and useful for planning interventions (Overton, 2000).

Purpose of the Study

As noted previously, behavioral descriptors for behavior rating scales are typically worded either positively or negatively. The informant is to determine whether the behavior is present (i.e., responding yes or no) or, using a rating scale, indicate intensity (e.g., rating on a scale of zero, behavior is not present, to four, behavior is always present). The purpose of this study was to determine if four behavior assessment

tools used positive or negative behavioral descriptors. Information presented includes test names, stated purposes, and recommended use of the tools (e.g., description, evaluation, screening); age levels of the targeted populations; and whether the behavioral descriptors are stated in positive, negative, or combination (positive and negative) terms. Two issues are addressed: (1) relativity of rater responses regarding whether items are worded as positive, negative, or a combination of both, and (2) the need for supporting evidence. Lastly, recommendations are given concerning the use of behavior rating scales for qualification for special education services.

Review of Behavioral Assessment Tools.

According to the assessment manuals, one behavior rating scale reviewed (*Conners Rating Scales System*) has behavioral descriptors that are all negatively worded, and one (*Walker-McConnell Scales of Social Competence & School Adjustment*) has behavioral descriptors that are all positively worded. Two of the four behavior rating scales use a combination of positive and negative behavioral descriptors (*Behavior Rating Profile & Child Behavior Checklist System*). In addition, two of the four are recommended to descriptive purposes (*Behavior Rating Profile & Child Behavior Checklist System*), one for evaluative purposes (*Behavior Rating Profile*), and three for screening purposes (*Child Behavior Checklist System, Conners Rating Scales System, & Walker/McConnell Scales of Social Competence and School Adjustment*).

Data contained in Table 1 reflect information gathered from the testing manuals. Included are recommended uses (description, evaluation, and/or screening) and behavioral descriptors of the behavior rating scales reviewed (positive, negative, or

combined positive and negative). Age levels of target populations and purposes of each scale follow:

1. Behavior Rating Profile (BRP): The BRP provides an ecological approach to behavior assessment in that it explores behaviors of children and adolescents (ages 6.6 through 18.6) across environments (Brown & Hammill, 1985). It can be used to identify or evaluate specific behaviors as well as plan academic or behavioral programs. There are six independent components available from which to choose (a teacher form, a parent form, three student forms, and a sociogram to be completed by peers). Thirty items are evaluated on a four point scale from "very much like the student" to "not at all like the student." Items reflect descriptions of behaviors and can be perceived as positive, negative, or neutral.
2. Child Behavior Checklist System (CBCL): The CBCL is a broad-spectrum behavior rating scale that provides descriptive data for the purposes of evaluation and intervention (Achenbach, 1992). The instrument is designed for use by parents and teachers to assess the competencies and problem behaviors of children and adolescents between the ages of four and eighteen. The 113 items are evaluated on a three point scale from zero (not true) to two (very true or often true). Items include descriptions of positive student characteristics (competencies), negative student characteristics, and neutral characteristics (i.e., allergies, asthma, shy, or timid).
3. Conners Rating Scales – Revised (CRS-R): Recommended uses of the CRS-R are screening, treatment monitoring, and as a direct clinical/diagnostic aid

(Conners, 1997). It consists of three scales (parent, teacher, and self-report) to diagnose and evaluate problem behaviors of children and adolescents ages three to seventeen (the self-report measure is recommended for use with ages twelve to seventeen). Twenty-eight items on the teacher scale and twenty-seven on the parent scale are evaluated on a four point scale from "not at all true" to "very much true." The items focus on sets of problems and are negatively worded (i.e., descriptive words such as *doesn't, fails to, lacks,* and *has difficulty*).

4. Walker/McConnell Scales of Social Competence and School Adjustment (SSCSA): The SSCSA are social skills rating scales that are intended for use by teachers and other school-based professionals to screen and identify social skills deficits among children and adolescents (Walker & McConnell, 1988). These scales focus on desired skills rather than problem-solving behaviors. There are two versions: an elementary version for use in grades K-6 and an adolescent version for use in grades 7-12. Forty-three items are evaluated on a five point scale from one (never) to five (frequently). The forty-three descriptors are positively worded and "written in a form appropriate for IEP objectives" (Walker & McConnell, 1988, pg. 2).

Initially, the primary author reviewed nine behavior rating scales and characterized the items as positive, negative, or a combination of positive and negative To corroborate these judgements, the secondary authors selected four instruments (described above) and independently rated each item as positive, negative, or neutral.

Inter-rater reliability was computed using the procedure, agreements divided by agreements and disagreements.

Results

Reliability coefficients for the four behavior rating scales ranged from .17 to .83. Evaluation of the Walker-McConnell Scale (r = .83) contained thirty-six agreements; all positively rated in the current evaluation. Disagreements consisted of seven items (e.g., "maintains eye contact when speaking or being spoken to") that were evaluated as neutral by one evaluator and positive by the other evaluator.

The Behavior Rating Profile-2 was characterized entirely negative by one evaluator (30 items) and predominantly negative (18 items) and "neutral" by the other for a reliability coefficient of .60. An example of an item that yielded disagreement was "daydreams." The Child Behavior Checklist yielded a reliability coefficient of .53. One evaluator considered the stems to be entirely negative. The other felt that fifty-three of the items were neutral or not obviously negative (e.g., "bites fingernails").

The Conners' produced the lowest reliability (.17). All items were considered to be negative by one evaluator. Twenty-three were considered, if not neutral, at least questionable (e.g., "poor in spelling," "lacks interest in school work," and "excitable, impulsive").

Discussion

Based on the evaluators' review, it appears as if most of the tools focus on the weaknesses of students with BD and do not present a well-rounded or complete picture of these students. It seems logical that assessment tools which focus primarily on negative

behaviors can influence people who work with these students to perceive them more negatively.

The fact that independent ratings of items as positive, negative, or neutral yielded inconsistent results is noteworthy. Teachers, parents, and others will probably demonstrate the same variability in their opinions about items. Unless informant information is treated cautiously, a range of perceptions can inaccurately characterize a student with an extremely negative label – behaviorally disordered.

Recommendations

This study corroborates the opinions of Witt, et. al. (1998) and the need for multiple informants and supplementary data collected through direct observation. The following recommendations are based on the instrument reviews:

1. Since rating scales require evaluative judgments by the informant they should be supplemented with directly observed behavioral data.
2. Rating scales can be used as part of an overall assessment process, however, educators should be aware that the scales reflect perceptions about the student. Multiple informants and inter-rater reliability checks can corroborate or contradict these perceptions of student behaviors.
3. Before utilizing a specific rating scale, educators should make sure it reflects the overall goal of the assessment process.
4. Care should be taken so that information about the student is not skewed toward the negative. Rating scales that are written in negative terms signal to the informant that a possible behavior problem exists.

5. Data on informant perceptions should be considered carefully. Patterns in a school district may suggest that certain classrooms and teachers yield higher incidences of students who are referred for services. Care in selecting instruments that give a fair appraisal of student behaviors, and policies that require corroboration of rating scale data are essential in making program decisions.

TABLE 1
Behavior Rating Scales

Tool	Description	Evaluation	Screening	-	+	-/+
Behavior Rating Profile	X	X				X
Child Beh. Checklist System	X		X			X
Conners Rating Scales Sys.			X	X		
Walker/McConnell Scales of Social Competence and School Adjustment			X		X	

- = Negative Terms; + = Positive Terms; -/+ = Positive and Negative Terms

References

Achenbach, T. M. (1992). Child Behavior Checklist. Burlington, VT: University of Vermont.

Bellack, A. S., & Hersen, M. (1988). Behavioral assessment: A practical handbook (3rd ed.). New York: Pergamon Press.

Brown, L. L., & Hammill, D. D. (1985). Behavior rating profile: An ecological approach to behavior assessment. Austin, TX: Pro-Ed.

Conners, C. K. (1997). Conners Rating Scales – Revised. North Tonawanda, NY: Multi-Health Systems, Inc.

Executive Committee of the Council for Children with Behavioral Disorders. (1989). White paper on best assessment practices for students with behavioral disorders: Accommodation to cultural diversity and individual differences. Behavioral Disorders, 14, 263-278.

Hallahan, D. P., & Kauffman, J. M. (1991). Exceptional children: Introduction to special education (5th ed.). Englewood Cliffs, N. J.: Prentice-Hall.

Mash, E. J., & Terdal, L. G. (Ed.). (1981). Behavioral assessment of childhood disorders. New York: Guilford Press.

Merrell, K. W. (1994). Assessment of behavioral, social, & emotional problems. New York: Longman.

Overton, T. (2000). Assessment in special education: An applied approach (3rd ed.). Upper Saddle River, NJ: Merrill.

Shapiro, E. S. (1987). Behavioral assessment in school psychology. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Swicegood, P. (1994). Portfolio-based assessment practices. Intervention in School and Clinic, 30(1), 6-15.

Walker, H. M., & McConnell, S. R. (1988). Walker/McConnell scales of social competence and school adjustment: A social skills rating scale for teachers. Austin, TX: Pro-Ed.

Witt, J. C., Elliott, S. N., Daly, E. J. III, Gresham, F. M., & Kramer, J. J. (1998). Assessment of at-risk and special needs children (2nd ed.). Boston: McGraw Hill.